



One-Photon Absorption Properties from a Hybrid Polarizable Density Embedding/Complex Polarization Propagator Approach for Polarizable Solutions

Hršak, Dalibor; Nørby, Morten Steen; Coriani, Sonia; Kongsted, Jacob

Published in:
Journal of Chemical Theory and Computation

Link to article, DOI:
[10.1021/acs.jctc.8b00155](https://doi.org/10.1021/acs.jctc.8b00155)

Publication date:
2018

Document Version
Peer reviewed version

[Link back to DTU Orbit](#)

Citation (APA):
Hršak, D., Nørby, M. S., Coriani, S., & Kongsted, J. (2018). One-Photon Absorption Properties from a Hybrid Polarizable Density Embedding/Complex Polarization Propagator Approach for Polarizable Solutions. *Journal of Chemical Theory and Computation*, 14(4), 2145-2154. <https://doi.org/10.1021/acs.jctc.8b00155>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Spectroscopy and Excited States

**One-photon absorption properties from a hybrid
Polarizable Density Embedding/Complex Polarization
Propagator approach for polarizable solutions**

Dalibor Hr̃scaron;ak, Morten Steen Nørby, Sonia Coriani, and Jacob Kongsted

J. Chem. Theory Comput., **Just Accepted Manuscript** • DOI: 10.1021/acs.jctc.8b00155 • Publication Date (Web): 16 Mar 2018

Downloaded from <http://pubs.acs.org> on March 22, 2018

Just Accepted

"Just Accepted" manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides "Just Accepted" as a service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. "Just Accepted" manuscripts appear in full in PDF format accompanied by an HTML abstract. "Just Accepted" manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are citable by the Digital Object Identifier (DOI®). "Just Accepted" is an optional service offered to authors. Therefore, the "Just Accepted" Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the "Just Accepted" Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these "Just Accepted" manuscripts.



ACS Publications

is published by the American Chemical Society, 1155 Sixteenth Street N.W., Washington, DC 20036

Published by American Chemical Society. Copyright © American Chemical Society. However, no copyright claim is made to original U.S. Government works, or works produced by employees of any Commonwealth realm Crown government in the course of their duties.

One-photon absorption properties from a hybrid Polarizable Density Embedding/Complex Polarization Propagator approach for polarizable solutions

Dalibor Hršak,[†] Morten Steen Nørby,[†] Sonia Coriani,[‡] and Jacob Kongsted^{*,†}

[†]*Department of Physics, Chemistry and Pharmacy, University of Southern Denmark,
Campusvej 55, 5230 Odense M, Denmark*

[‡]*Department of Chemistry, Technical University of Denmark, Kemitorvet Build. 207, 2800
Kgs. Lyngby, Denmark*

E-mail: kongsted@sdu.dk

Abstract

We present a formulation of the polarizable density embedding (PDE) method in combination with the complex polarization propagator (CPP) method for the calculation of absorption spectra of molecules in solutions. The method is particularly useful for the calculation of near-edge X-ray absorption fine structure (NEXAFS) spectra. We compare the performance of PDE-CPP with the previously formulated polarizable embedding (PE) CPP model for the calculation of the NEXAFS spectra of adenine, formamide, glycine and adenosine triphosphate (ATP) in water at the carbon and nitrogen K-edges, as well as of formamide and glycine at the oxygen K-edge. In general, we find only minor differences between the performance of PDE and PE for the targeted parts of the spectra, except in the case of transitions involving Rydberg states,

for which non-electrostatic effects are found to be important.

Introduction

Multiscale techniques have found a significant place in the domain of molecular modeling due to their efficiency and versatility in tackling the challenges of predicting properties of large molecular systems in comparison to full quantum-mechanical descriptions.¹ A popular set of multiscale methods are hybrid quantum-mechanical/molecular-mechanical (QM/MM) methods that focus the computational effort on the chemically important part of the system while addressing the rest of the system in a less expensive fashion using classical force fields.²⁻⁵ Among the plethora of QM/MM type multiscale methods^{4,6,7} are the polarizable embedding^{8,9} (PE) and the QM/QM/MM polarizable density embedding¹⁰ (PDE) models, both developed in recent years to facilitate efficient and accurate calculation of environment effects on molecular response properties. Standard response theory¹¹⁻¹³ has found its application in predicting various important response properties, such as molecular polarizabilities, UV/vis spectra, multiphoton absorption, etc.¹⁴ These properties are calculated from the linear or higher order response functions and their residues. Standard response theory assumes an infinite lifetime of the excited states. The resulting response functions contain poles on the frequency axis at positions corresponding to resonance energies of the system. On one side, this yields a recipe to compute, *e.g.*, one-photon absorption spectra, from the poles (excitation energies) and residues (oscillator strengths) of the dipole-dipole linear response function. On the other side, due to their pole structure, the response functions themselves (as well as some of their residues) become unphysical in near-resonance and resonance regions of the spectrum and do not represent the observable of interest. Moreover, even if computing absorption spectra from individual excitation energies and strengths is not difficult in cases where the excited states are well resolved (*i.e.*, where there is low density of states in the observed region), it may become quite challenging when targeting parts of the spectrum

where the density of states is high. This is for instance the case for large molecular systems, and for the X-ray frequency region.

One way to overcome the above-mentioned problems is by taking into account the finite lifetimes of excited states through formulations of damped response theory,^{15–18} as done *e.g.* within the complex polarization propagator^{15,16,19} (CPP) method, where the response functions are well defined on the entire spectrum. Damped response theory has already been successfully implemented in combination with the PE model for both density functional theory²⁰ (DFT) and coupled cluster (CC).²¹ In PE, the environment is accounted for in an explicit manner through a set of multipole moments and polarizabilities typically with a local origin on the atoms defining the environment. The multipole moments give rise to electrostatic interaction terms, while polarizabilities allow for inclusion of the environment polarization terms when calculating both the ground state properties and molecular response properties. Thus, the changes in the molecular electron density upon *e.g.* an electronic excitation are reflected in the environment. It has previously been observed that PE lacks a fully satisfactory description of intermolecular interactions in the near vicinity of the chromophore, where the wave function of the chromophore begins to overlap with the wave function of the environment,²² meaning that non-electrostatic repulsion can play an important role.²³ The lack of the non-electrostatic repulsion may result in an artificial relaxation of the chromophore electron density by extending it unnaturally into the environment.²⁴ This problem has been solved by formulating the PDE model in which the environment is not described through atom-centered multipoles, but through a set of fragment charge densities calculated for isolated fragment molecules by using a QM method, typically Hartree-Fock or DFT. This approach can be combined with PE, where only the nearest solvent molecules are described with PDE, thus resulting in a fully polarizable three-layered QM/QM/MM model. This model is capable of describing short-range non-electrostatic repulsion effects and thereby preventing density leak from the chromophore into its environment. It has been shown that such wave function confinement in PDE can play an important role in modeling

solvent effects on electronic transitions.^{24,25}

The goal of this paper is to explicitly formulate the PDE-CPP model within TDDFT theory and compare its performance with the already reported PE-CPP model. For this PE-CPP and PDE-CPP have been applied to the calculations of the near-edge X-ray absorption fine structure spectra^{26,27} (NEXAFS) of adenine, formamide, zwitterionic glycine and adenosine triphosphate (ATP) in aqueous solutions. Where available, our calculated spectra are compared with the experimental results.

Theory

The PDE model

The PDE model, as introduced in Ref. 10, partitions a given system into three layers with different levels of description: *i*) the “core” region, *ii*) the “inner” region and *iii*) the “outer” region. Since we will discuss core-electron spectra as exemplificative applications of our methodology, we will from now on refer to the “core” region as “active quantum” region, to avoid any confusion in the use of the word core. The active quantum region is used to describe the solute molecule in a fully quantum mechanical fashion, whereas the inner region is given as a set of fragment charge densities calculated at a self-consistent-field (SCF) level for isolated solvent molecules. The outer region is described classically, as a set of multipoles (partial charges, dipoles, quadrupoles, and so on) centered at the atoms defining the environment. The atomic centers in both the inner and outer regions are also assigned with anisotropic dipole-dipole polarizabilities. The multipoles and polarizabilities are obtained through QM calculations of local properties on isolated solvent molecules. The embedding potential that describes the inner region is denoted as FDP y R, where FD denotes the description of electrostatic interactions with fragment electron densities, y denotes the type of the polarizabilities ($y=1$ indicates isotropic and $y=2$ indicates anisotropic polarizabilities) and R denotes the inclusion of the non-electrostatic repulsion in the model. The correspond-

ing embedding potential for the outer region is denoted as $MxPy$, where x specifies the order in multipoles and y again indicates the type of the polarizabilities, in the same way as for the inner region.

The interactions between the solute in the active quantum region and solvent molecules in the inner and outer regions are accounted for through one-electron effective environment operators. The first term is the fragment density operator, that describes the electrostatic interactions between the active-quantum electron density and the inner region's nuclei and electrons. A matrix element in an atomic orbital (AO) basis of this operator is given as

$$v_{\mu\nu}^{\text{fd}} = - \sum_{a=1}^{N_{\text{in}}} \sum_{m \in a}^{M_a} \left\langle \chi_{\mu}(\mathbf{r}) \left| \frac{1}{|\mathbf{R}_m - \mathbf{r}|} \right| \chi_{\nu}(\mathbf{r}) \right\rangle Z_m + \sum_{a=1}^{N_{\text{in}}} \sum_{\gamma \delta \in a} \left\langle \chi_{\mu}(\mathbf{r}) \chi_{\gamma}(\mathbf{r}') \left| \frac{1}{|\mathbf{r}' - \mathbf{r}|} \right| \chi_{\nu}(\mathbf{r}) \chi_{\delta}(\mathbf{r}') \right\rangle D_{\gamma\delta}^{\text{fd}} \quad , \quad (1)$$

where N_{in} is the number of inner region fragments and M_a is the number of nuclei in fragment a , Z_m is the charge of the nucleus m in fragment a of the inner region. χ_{μ} and χ_{ν} denote active-quantum basis functions and χ_{γ} and χ_{δ} denote basis functions of an inner region fragment a , while $D_{\gamma\delta}^{\text{fd}}$ is an element of the fragment charge density matrix. Since the inner region encompasses the solvent molecules in the nearest vicinity of the solute, short-range interactions become significant due to wave-function overlap between solute and solvent. Such overlap effects are introduced through the non-electrostatic repulsion operator constructed using the Huzinaga-Cantu projection method,²⁸ which simulates orthogonality between the orbitals in the active-quantum and inner regions. A matrix element of this operator has the form

$$v_{\mu\nu}^{\text{rep}} = - \sum_{a=1}^{N_{\text{in}}} \sum_{\gamma \delta \in a} W_{\gamma\delta} S_{\mu\gamma} S_{\nu\delta} \quad , \quad (2)$$

where $W_{\gamma\delta}$ is the energy-weighted density matrix containing the MO energies of the inner region fragment a and the corresponding MO coefficients. $S_{\mu\gamma}$ and $S_{\nu\delta}$ are elements of the overlap matrix between the orbitals in the active quantum region and the inner region.

The active-quantum electron density also experiences electrostatic interactions with the multipoles in the outer region. The matrix element of this electrostatic multipole operator is given as

$$\begin{aligned}
 v_{\mu\nu}^{\text{mul}} &= \sum_{b=1}^{N_{\text{out}}} \sum_{s \in b}^{S_b} \sum_{|k|=0}^K \frac{(-1)^{|k|}}{k!} \left\langle \chi_{\mu}(\mathbf{r}) \left| \nabla_{\mathbf{r}}^k \frac{1}{|\mathbf{R}_s - \mathbf{r}|} \right| \chi_{\nu}(\mathbf{r}) \right\rangle M_s^{(k)} \\
 &= \sum_{b=1}^{N_{\text{out}}} \sum_{s \in b}^{S_b} \sum_{|k|=0}^K \frac{(-1)^{|k|}}{k!} t_{\mu\nu}^{(k),s} M_s^{(k)} \quad .
 \end{aligned} \tag{3}$$

In this equation, the index b runs over all of the fragments in the outer region, $M_s^{(k)}$ is the k^{th} order multipole located at atom s belonging to fragment b and $t^{(k),s}$ is the k^{th} order electric field integral at the site s . Furthermore, we have introduced a compact notation where k is a multi-index.^{29,30}

Polarization effects are treated on the same footing for both the inner and the outer regions, *i.e.* through anisotropic dipole-dipole polarizabilities assigned to atom centers of the molecules in the solvent. A matrix element of the polarization operator is given as

$$\begin{aligned}
 v_{\mu\nu}^{\text{ind}} &= - \sum_{a=1}^{N_{\text{in}}+N_{\text{out}}} \sum_{s \in a}^{S_a} \left\langle \chi_{\mu}(\mathbf{r}) \left| \nabla_{\mathbf{r}} \frac{1}{|\mathbf{R}_s - \mathbf{r}|} \right| \chi_{\nu}(\mathbf{r}) \right\rangle \boldsymbol{\mu}^{\text{ind},s} \\
 &= - \sum_{a=1}^{N_{\text{in}}+N_{\text{out}}} \sum_{s \in a}^{S_a} t_{\mu\nu}^{(1),s} \boldsymbol{\mu}^{\text{ind},s} \quad .
 \end{aligned} \tag{4}$$

The physical interpretation of the polarization operator is that the active-quantum electron density interacts with the induced dipole moment $\boldsymbol{\mu}^{\text{ind},s}$ in the surrounding through the electric field. The dipole moment at site s is here induced by the electric field of the electron density in the active-quantum region, in addition to the active-quantum nuclei, the fragment charge densities and nuclei in the inner region and the multipoles in the outer region, except those belonging to the same fragment a . The site also experiences the field arising from all of the other induced dipole moments, which results in the following expression for the induced

dipole moments

$$\boldsymbol{\mu}^{\text{ind}} = \boldsymbol{\alpha}(\mathbf{F}^{\text{el}} + \mathbf{F}^{\text{nuc}} + \mathbf{F}^{\text{fd}} + \mathbf{F}^{\text{mul}} + \mathbf{F}^{\text{ind}}) \quad (5)$$

The expressions for the above given fields can be found in Refs. 8–10. The interplay of mutual polarizations of the classical sites in the inner and outer regions with the active-quantum region is solved efficiently through the construction of a classical response matrix³¹ that contains all the polarizabilities and the dipole-dipole electrostatic interaction tensors.

The complex polarization propagator in PDE context

For the formulation of the PDE-CPP approach we rely on density functional theory (DFT). As such, we adopt an exponential parameterization of the time-dependent wave function¹¹

$$|\psi(t)\rangle = \exp(i\hat{\boldsymbol{\kappa}}(t)) |0\rangle \quad , \quad (6)$$

where $|0\rangle$ is a SCF reference wave function and $\hat{\boldsymbol{\kappa}}(t)$ is a time-dependent orbital rotation operator that generates variations of the reference state in the orbital space. The time-dependent parameters in the latter operator are expanded in orders of the perturbation as

$$\boldsymbol{\kappa}(t) = \boldsymbol{\kappa}^{(1)}(t) + \boldsymbol{\kappa}^{(2)}(t) + \dots \quad (7)$$

The CPP response functions are derived based on the Ehrenfest’s theorem and the Liouville equation on the time-variation of the expectation value of an operator.¹² The final form of the linear response function is¹¹

$$\left\langle \left\langle \hat{A}; \hat{V} \right\rangle \right\rangle_{\omega}^{\gamma} = -i\mathbf{A}^{\dagger} \boldsymbol{\kappa}^{\omega} \quad (8)$$

where ω is the frequency of the external perturbation and \mathbf{A} is the property gradient for the operator \hat{A} . In order to calculate the linear response function, one first has to find the $\boldsymbol{\kappa}^{\omega}$

parameters, by solving the complex linear equation¹⁹

$$\boldsymbol{\kappa}^\omega = i[\mathbf{E} - (\omega + i\gamma)\mathbf{S}]^{-1}\mathbf{V}^\omega \quad (9)$$

In this equation, the resulting matrices are the Hessian matrix \mathbf{E} , generalized overlap matrix \mathbf{S} , and the property gradient \mathbf{V}^ω of the external perturbation \hat{V}^ω . The relaxation parameter γ is a constant that enables to define real and imaginary parts of the response function at a resonance frequency and is connected to the lifetime of the corresponding excited state.

In the PDE-CPP context, the interaction of the active-quantum region with the environment enters the Hessian matrix \mathbf{E} and adds the PDE term to the vacuum contribution: $\mathbf{E} = \mathbf{E}^{\text{vac}} + \mathbf{E}^{\text{PDE}}$. The embedding operator \hat{v}^{PDE} contains the sum of the contributions given in eqs. (1) to (4). The PDE-Hessian linearly transformed vector has the following form

$$\begin{aligned} \mathbf{E}^{\text{PDE}}\boldsymbol{\kappa}^\omega &= -\langle 0 | [\hat{\mathbf{q}}, [\hat{\boldsymbol{\kappa}}^\omega, \hat{v}^{(0),\text{PDE}}] + \hat{v}^{\omega,\text{PDE}}] | 0 \rangle = -\langle 0 | [\hat{\mathbf{q}}, \hat{Q}_1^\omega + \hat{Q}_2^\omega] | 0 \rangle \\ \hat{Q}_1^\omega &= [\hat{\boldsymbol{\kappa}}^\omega, \hat{v}^{(0),\text{PDE}}] = \tilde{V}^{\text{fd}}(\boldsymbol{\kappa}^\omega) + \tilde{V}^{\text{rep}}(\boldsymbol{\kappa}^\omega) + \tilde{V}^{\text{mul}}(\boldsymbol{\kappa}^\omega) - \boldsymbol{\mu}^{\text{ind}} \tilde{F}^{\text{el}}(\boldsymbol{\kappa}^\omega) \\ \hat{Q}_2^\omega &= \hat{v}^{\omega,\text{PDE}} = -\tilde{\boldsymbol{\mu}}^{\text{ind}}(\boldsymbol{\kappa}^\omega) \hat{F}^{\text{el}} \end{aligned} \quad (10)$$

The terms \tilde{V}^{fd} , \tilde{V}^{rep} , \tilde{V}^{mul} , \tilde{F}^{el} and $\tilde{\boldsymbol{\mu}}^{\text{ind}}$ are one-index transformed interaction operators. We see that the only difference between the PDE- and the PE-CPP model is in the \hat{Q}_1^ω term, where the additional terms occur that describe the interaction between the active-quantum and the inner region. Furthermore, the induced dipole moments in \hat{Q}_1^ω and \hat{Q}_2^ω are also affected by the fields arising from the fragment densities terms. A density-driven polarization routine is retained in this model, whereby the active-quantum region's electron density produces the electric field, that induces the electric dipoles in the environment, which in turn modify the embedding operator that enters the response equations.³²

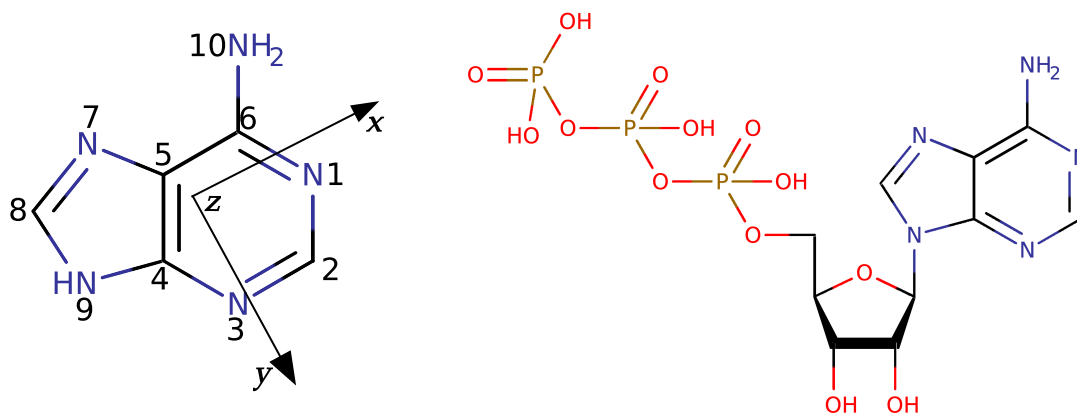


Figure 1: Structure of the adenine molecule (with heavy atom numbering and Cartesian axes) and structure of the ATP molecule. Note that the z -axis is perpendicular to the molecular plane. The atom numbering for the adenine part of ATP is retained from the isolated adenine molecule. Atom **N1** is the protonation site for the case of protonated adenine and ATP. The images were generated using MarvinSketch.³³

Computational details

Fig. 1 depicts adenine and ATP. The protonated forms of adenine and ATP include an additional proton attached to the **N1** atom. The geometries of adenine and ATP in gas phase, as well as their protonated forms were optimized using DFT with the M06-2X functional³⁴ and the 6-31+G* basis set^{35–38} using the Gaussian package.³⁹

For the preparation of water-solvated adenine and ATP, we used the Maestro package.⁴⁰ We placed either adenine or ATP molecules into water boxes containing ~ 500 water molecules, the buffer-size thereof being 10 Å. For the protonated forms of adenine and ATP, an additional Cl^- ion was placed in the box in order to neutralize the system. Thereafter, a molecular dynamics (MD) simulations 10 ns long were performed with the Desmond package,⁴¹ adopting the same settings as in our previous work (see Ref. 42). From the simulations, snapshots in 20 ps time intervals were extracted for further analysis. Adenine or ATP molecules (either neutral or protonated) in these snapshots were QM/MM geometry optimized with M06-2X/6-31+G*/OPLS2005⁴³ using the Qsite program.^{44–46} In these QM/MM optimizations, the QM region consisted of either adenine or ATP, while all water molecules

(and Cl^- ion where appropriate) were described with the above given force-field and were furthermore kept frozen. For formamide we took geometries produced in a rigid-body MD simulation, the details of which can be found in Ref. 47. We included the solvent molecules in a solvent sphere of a buffer size of 12 Å. For glycine, geometries of the molecule in a box consisting of 512 water molecules were taken from a rigid-body MD simulation as previously reported.⁴⁸

The embedding potentials FDP2R and M2P2 describing the environment used in PDE-CPP and PE-CPP calculations, respectively, were obtained using the PE Assistant Script²⁹ (PEAS) and with the Dalton package for the calculation of fragment densities, and with Molcas⁴⁹ using the LoProp procedure⁵⁰ for the calculation of local properties (multipoles and polarizabilities).

The X-ray absorption spectra of all systems were calculated using the fully long-range corrected version of CAM-B3LYP^{51,52} in combination with the aug-cc-pVDZ basis set,^{53–56} and the program used was a local version of Dalton 2016.⁵⁷ For the solvated systems we used 50 selected snapshots of adenine, 120 snapshots of formamide and glycine and 10 snapshots of ATP. Absorption spectra of the observed systems were obtained from the damped linear response functions as the imaginary parts of the molecular polarizabilities (α), calculated at frequencies corresponding to the near carbon, nitrogen or oxygen K-edge parts of the spectra. The broadening parameter γ was set equal to 1000 cm^{-1} (in practice, this corresponds to the half-width-at-half-maximum of an unnormalized Lorentzian line shape function). The spectra have been smoothed by interpolation with a cubic spline and the absorption cross sections were calculated using the formula⁵⁸

$$\sigma(\omega) = \frac{\omega}{\varepsilon_0 c} \text{Im}\{\bar{\alpha}(\omega)\}; \quad \bar{\alpha}(\omega) = \frac{1}{3}[\alpha_{xx}(\omega) + \alpha_{yy}(\omega) + \alpha_{zz}(\omega)], \quad (11)$$

where ω is the frequency, ε_0 is the electric constant and c is the speed of light.

Results

Adenine in vacuo The CPP spectra obtained for the optimized structure of adenine in vacuum, and for the optimized adenine snapshots in water with PDE or PE, can be found in Fig. 2. A compilation of the corresponding positions of the peak maxima and their assignments can be found in Table 1. The CPP results reproduce well the main features of

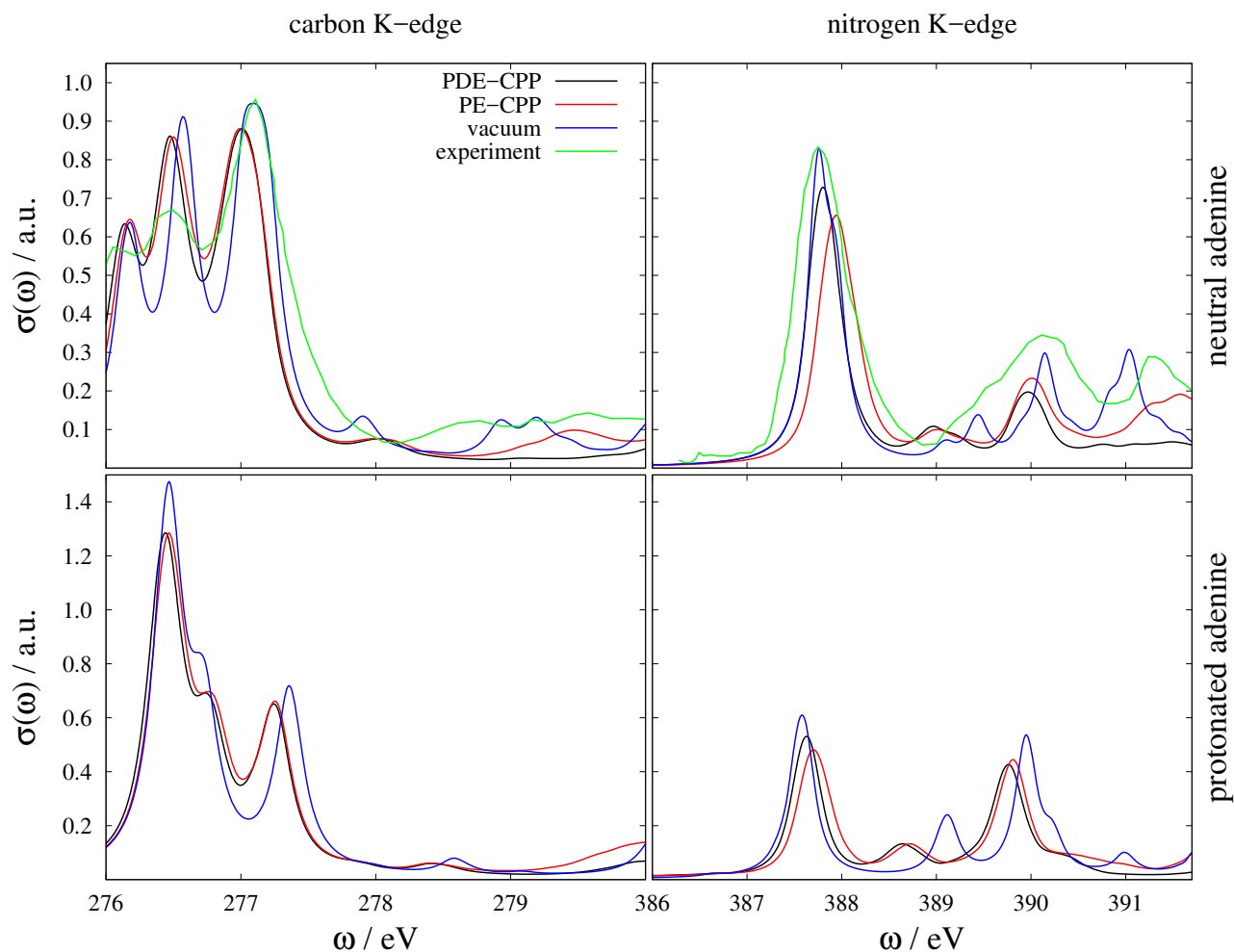


Figure 2: Calculated X-ray spectra of gas phase and aqueous adenine in its neutral or protonated states. The experimental results from Ref. 59 corresponding to adenine in the gas phase are also shown. The calculated solvated spectra were obtained as averages based on 50 selected snapshots, see text for details. The experimental results have been shifted (by -10.28 eV for the carbon K-edge and by -11.72 eV for the nitrogen K-edge) and scaled to be comparable with the calculated results, and they were digitized using WebPlotDigitizer.⁶⁰

the experimental spectra from Ref. 59. The computed peaks are found at ~ 10 eV – 12 eV

lower energies than the corresponding experimental peaks. The Carbon K-edge part of the simulated spectrum of neutral adenine is characterized by three consecutive peaks at 276.2 eV, 276.6 eV and at 277.1 eV, which correspond to the three experimental peaks at 286.4 eV, 286.8 eV and 287.2 eV. The peak at 276.2 eV is polarized along the z -axis and arises from excitation of an electron from the $1s$ -orbital localized on the **C2** atom. The 276.6 eV peak is also z -polarized and arises from a core excitation from the **C8** atom, while the peak at 277.1 eV arises from the **C6** atom. In the protonated spectrum there are only two distinctive peaks, one at 276.5 eV and other at 277.4 eV, where the former is a **C2** core excitation and the latter is a of **C6** core excitation. The **C2** peak also has a shoulder on the right-hand side, that is a mixture of **C4** and **C8** excitations.

The Nitrogen K-edge spectrum of the neutral adenine is characterized by one intense peak at 387.7 eV (corresponding to the experimental peak at 399.5 eV) and a series of four weaker peaks. The intense peak arises from excitation from $1s$ -type MOs at the **N7** and **N3** atoms for its z -component of the dipole transition moment. The small peak at 389.4 eV is also z -polarized and is characterized by excitation from the amino nitrogen, **N10**. In the experiment, this peak is not distinguishable, because it is overlapping with the more intense peak. The x -, y - and z -components of the dipole transition moment associated with the peak at 390.1 eV are composed principally from s -excitations at atoms **N1**, **N3** and **N9** respectively, as the peak is not polarized exclusively along one axis. This peak corresponds to the experimental peak at 401.9 eV. For the peak at 391.0 eV, the x - and y -components of the dipole transition moment arise from atom **N10**, whereas the z -component has its origin in the excitation from the **N7** atom, although it has a lower contribution. This last peak corresponds to the experimental peak at 403.1 eV and it is strongly attenuated in the protonated vacuum spectrum.

The peak assignments given here are consistent with the assignments given by Plekan *et al.*⁵⁹ obtained from ADC(2) stick spectra calculations. Since the CPP formalism directly computes the total cross section, and not the individual electronic transitions underlying each

given band, information about the nature of the transitions is obtained from an analysis of the composition of the response vectors entering the computational expression of the total cross section (isotropic imaginary polarizability). Thus, unlike ADC(2), the CPP formalism does not directly provide information on virtual orbitals involved in the excitations.

Aqueous adenine The spectra of neutral aqueous adenine are also characterized by three distinct peaks at the near carbon K-edge, and there is very little difference between the performance of the PDE- and PE-models. The peaks located at 276.1 eV and 276.2 eV (for PDE and PE, respectively) belong to a pure transition from the 1s-orbital at atom **C2**, while the peak at 276.5 eV originates from the **C8** 1s-orbital. The 277.0 eV peak arises from the **C6** atom and it is also “red-shifted” (*i.e.*, at a lower frequency) as compared to the vacuum case. Unlike the vacuum case for the protonated adenine, the shoulder of the most intense peak of protonated aqueous adenine appears as a distinct peak with its origins in a **C4** excitation.

The nitrogen K-edge part of the spectrum is characterized by one intense peak and two peaks with less intensity. The intense peak is composed of excitations from **N7** and **N3** atoms for PDE (387.8 eV) and of excitations from **N7** and **N1** for PE (387.9 eV), where the PDE peak is more intense than the PE peak. The next peak is located at 389.0 eV for both PDE and PE, it has low intensity and is broad compared to other analyzed peaks. It is of the same composition for both embedding models, namely a 1s excitation from the amino nitrogen atom **N10**. This peak corresponds to the vacuum peak at 389.4 eV and it is strongly affected by the hydrogen bonding of the amino-nitrogen with the solvent molecules.

The third peak is at 390.0 eV for PDE and at 390.1 eV for PE and it arises principally from the **N9** excitation. The fourth vacuum peak at 391.0 eV is strongly attenuated in the aqueous solution, especially for PDE. The strong solvent effects and the difference between PDE and PE confirms the observation from Plekan *et al.*⁵⁹ that this is a Rydberg transition from the amino-nitrogen. Because the Rydberg states are usually strongly affected by the

Table 1: Positions of peaks (in eV) and their assignments in the carbon and nitrogen K-edges of the X-ray absorption spectra of adenine and ATP in vacuum and in aqueous solution, all in both neutral and protonated forms calculated with CPP, PDE-CPP or PE-CPP.

	Adenine		Adenosine-triphosphate	
	vacuum	PDE	vacuum	PE
neutral	C K-edge	276.2 (C2)	276.2 (C2)	276.4 (C2,C5,C8)
			276.3 (C8)	
		276.5 (C8)	276.6 (C5)	
		277.0 (C6)	277.0 (C6)	277.0 (C6)
neutral	N K-edge	387.7 (N7,N3)	387.9 (N1)	387.8 (N3,N7,N1)
		389.4 (N10)	389.3 (N10)	389.1 (N10)
		390.1 (N1,N3,N9)	389.9 (N9)	390.0 (N9)
		391.0 (N10,N7)	391.2 (N10)	391.3 (N10)
protonated	C K-edge	276.5 (C2)	276.2 (C8)	276.4 (C8,C2)
		276.7 (C4,C8)	276.5 (C2)	276.6 (C4)
		277.4 (C6)	277.4 (C6)	277.3 (C6)
protonated	N K-edge	387.6 (N7)	387.6 (N3)	387.7 (N7, N3)
		389.1 (N10)	389.2 (N10)	388.7 (N10)
			389.7 (N9)	
		390.0 (N9)	390.0 (N1)	389.9 (N9,N1)
		391.0 (N10)	390.9 (N10,N7)	389.9 (N9,N1)

solvent through non-electrostatic repulsion,^{61,62} it is expected that the non-electrostatic repulsion contribution in PDE will additionally destabilize the transition, thereby producing the PDE/PE difference. The protonated PDE and PE spectra show pronounced solvent effects on the **N10** and **N9** peaks, where these peaks are red shifted with respect to the vacuum case. Unlike PE, the PDE **N9** peak also contains a significant contribution from the **N1** excitation, but this contribution does not produce a visible difference in the position and shape of the peak from the PE case.

Aqueous formamide The next system studied is formamide. Formamide is an interesting test case since it is well known that the lower lying part of the electronic absorption spectrum of formamide contains a mixture of localized valence and Rydberg transitions. It can be expected that PDE and PE describe such Rydberg transitions differently, so it is of interest to investigate if this also impacts calculations of the X-ray absorption spectra. The calculated NEXAFS spectra of formamide at its carbon, nitrogen and oxygen K-edges are shown in Fig. 3 and in comparison with the experiment for its carbon and oxygen K-edges, whereas we didn't find experimental results for nitrogen K-edge in the literature. Formamide is a small and relatively simple molecule, with no mixed excitations from several heavy atoms of the same element, making it easy to interpret. The carbon K-edge has a single intense transition, of $C1s \rightarrow \pi^*$ character, where the differences between PDE and PE are negligible. However, the nitrogen and oxygen K-edge absorption spectra are visibly different for PDE and PE. The nitrogen K-edge has the dominant $N1s \rightarrow \pi^*$ transition at a higher energy and with lower intensity for PDE than for PE. This transition from the amino-group corresponds to the second peak in adenine, but the effect of the non-electrostatic repulsion is more visible for formamide. The "blue shift" of PDE as compared to PE can be explained by the virtual π^* orbital confinement and its destabilization due to non-electrostatic repulsion in PDE. Higher excitation at the nitrogen K-edge are even more different for the two embedding models, because they are probably of Rydberg character, as observed in gas phase.⁶⁴ For the

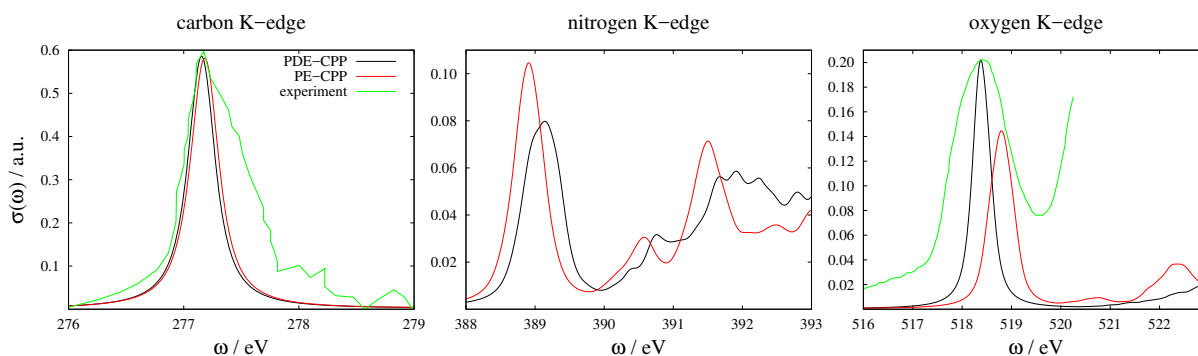


Figure 3: Calculated X-ray absorption spectra of gas phase and aqueous formamide at its carbon, nitrogen and oxygen K-edges along with the experimental results where available (in arbitrary units on the y -axis) from Ref. 63. The experimental spectra were digitized using WebPlotDigitizer⁶⁰ and were shifted by -11 eV for carbon and by -14.2 eV for oxygen K-edge and scaled to overlap with the PDE spectra. Higher transitions at the oxygen K-edge cannot be experimentally validated because of solvent absorption in the experiment.

$O1s \rightarrow \pi^*$ transition the effect of the orbital confinement has an inverse effect, *i.e.* the PDE peak is located at a lower frequency and it is more intense than for PE. It has previously been observed experimentally that a water solvent does not have much influence on the X-ray spectrum of formamide at its carbon K-edge, whereas it does at the oxygen K-edge.⁶³ Along this line, we here observe a greater difference in the two solvation models exactly for the oxygen K-edge compared to the carbon K-edge.

Aqueous zwitterionic glycine The next system investigated is glycine in water solution. For glycine an experimental NEXAFS spectrum in water has been measured⁶⁵ and thereby provides a good benchmark for comparison of PDE and PE. The NEXAFS spectra of aqueous glycine at its carbon, nitrogen and oxygen K-edges are shown in Fig. 4 in comparison with the experimental results. The carbon K-edge is characterized by an intense peak corresponding to a transition from the carboxylic C-atom, *i.e.* a $C1s \rightarrow \pi^*$ transition. For that peak, there is no difference between the performance of PDE and PE. The following series of less intense peaks are in Ref. 65 described as $C1s \rightarrow \sigma^*$ /Rydberg transitions. In this part of the spectrum, we observe a greater difference in the performance between PDE and PE. This is expected because these Rydberg states are diffuse in nature and more sensitive to the effects

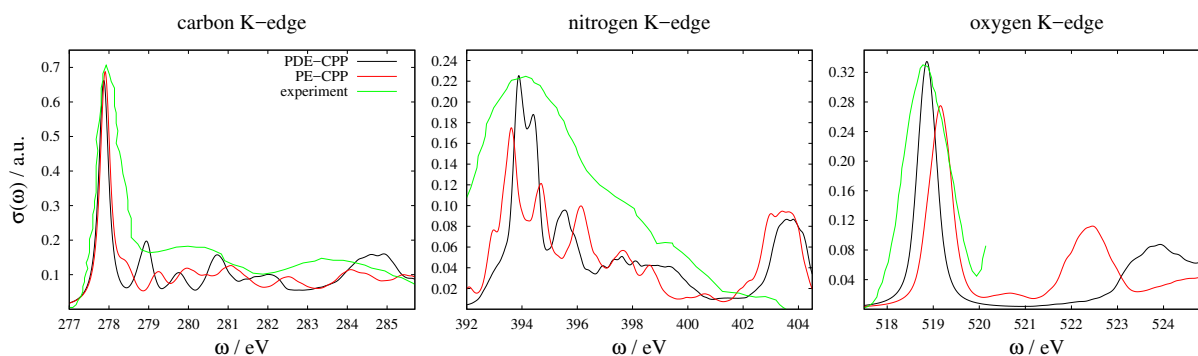


Figure 4: NEXAFS spectra of aqueous glycine calculated at its carbon, nitrogen and oxygen K-edges with either PDE-CPP or PE-CPP and compared with the experimental spectra (in arbitrary units on the y -axis) from Ref. 65. The experimental spectra were shifted by -10.6 eV, -12.5 eV and -13.4 eV for carbon, nitrogen and oxygen K-edges, respectively, and scaled to overlap with the PDE spectra. The experimental spectra were digitized using WebPlotDigitizer.⁶⁰ Higher transitions at the oxygen K-edge cannot be experimentally validated because of solvent absorption in the experiment.

of the nearest environment that is more accurately described with PDE.

The nitrogen K-edge spectrum has a series of overlapping transitions originating from the amino-group. From the experimental data it is seen that these transitions are merged into one wide band without distinct peaks. The main PDE peak is located at a higher frequency than the PE peak, which is similar to what is observed for formamide but opposite from adenine. The oxygen K-edge also shows opposite tendencies from the nitrogen K-edge, *i.e.* the PDE peak lies at a lower frequency compared to the PE peak. The reason for this observation is due to a different orientation of the nearest water molecules and thus an opposite direction of the hydrogen bonds between the solutes and the solvent. The amino-group is a hydrogen donor and oxygen is a hydrogen acceptor. For PE, the lack of non-electrostatic repulsion causes the electron density to be more extended into the environment for hydrogen acceptors whereas for PDE the electron density is more confined at the molecule. For hydrogen donor groups, such as the amino-group, the negative partial charge of the nearest solvent oxygen will repel any excess charge in PE. These effects will result in opposite shifts between PDE and PE at the nitrogen and oxygen K-edges in both formamide and glycine.

ATP in vacuo As the last system studied we here consider a much more complex case of ATP. The geometry of a single structure of isolated ATP was optimized, and its NEXAFS spectrum was calculated in the same manner as for adenine in vacuum. The spectrum is shown in Fig. 5. The carbon K-edge part of the spectrum of neutral ATP reveals that there are two pronounced peaks in the absorption region of the adenine carbons, instead of three peaks as in isolated adenine. The left-hand side intense peak is in fact composed of three overlapping peaks at 276.2 eV, 276.3 eV and 276.6 eV with their origins in the excitations of the 1s-electrons from **C2**, **C8** and **C5**, respectively. The other intense peak at 277.0 eV originates from **C6**. In the region between 277.9 eV and 279.4 eV, a series of five peaks is located, all originating from the excitation of the 1s-electrons at the ribose carbon atoms. The nitrogen K-edge part of the X-ray spectrum contains four distinctive peaks as in the case of adenine, where the most intense is located at 387.9 eV, stemming from the **N1** atom. The peak from the amino **N10** atom is located at 389.3 eV, while at 389.9 eV one can find the peak arising from the **N9** atom. Finally, the peak at 391.2 eV also originates from the amino group, as in the case of adenine.

In the carbon K-edge spectrum of the protonated form, the **C8** peak is shifted to 276.2 eV and the **C2** peak is shifted to 276.5 eV, while the **C5** peak is indistinguishable. The **C6** peak is shifted to 277.4 eV and it has a lower intensity than in the neutral form. In the nitrogen K-edge spectrum, the first intense peak is weaker than in the neutral form and it corresponds to **N3** transition. The following series of three consecutive peaks at 389.2 eV, 389.7 eV and 390.0 eV are **N10**, **N9** and **N1** core excitations, while the peak at 391.0 eV is weaker than in the neutral case.

Aqueous ATP The carbon K-edge of the spectrum for aqueous neutral ATP has two distinctive peaks (276.4 eV and 277.0 eV) corresponding to the experimental peaks at 286.5 eV and 287.3 eV in Ref. 66. The difference between the adenine and the ATP spectra, namely three versus two peaks, was also found in the experiment. The peak at 276.4 eV

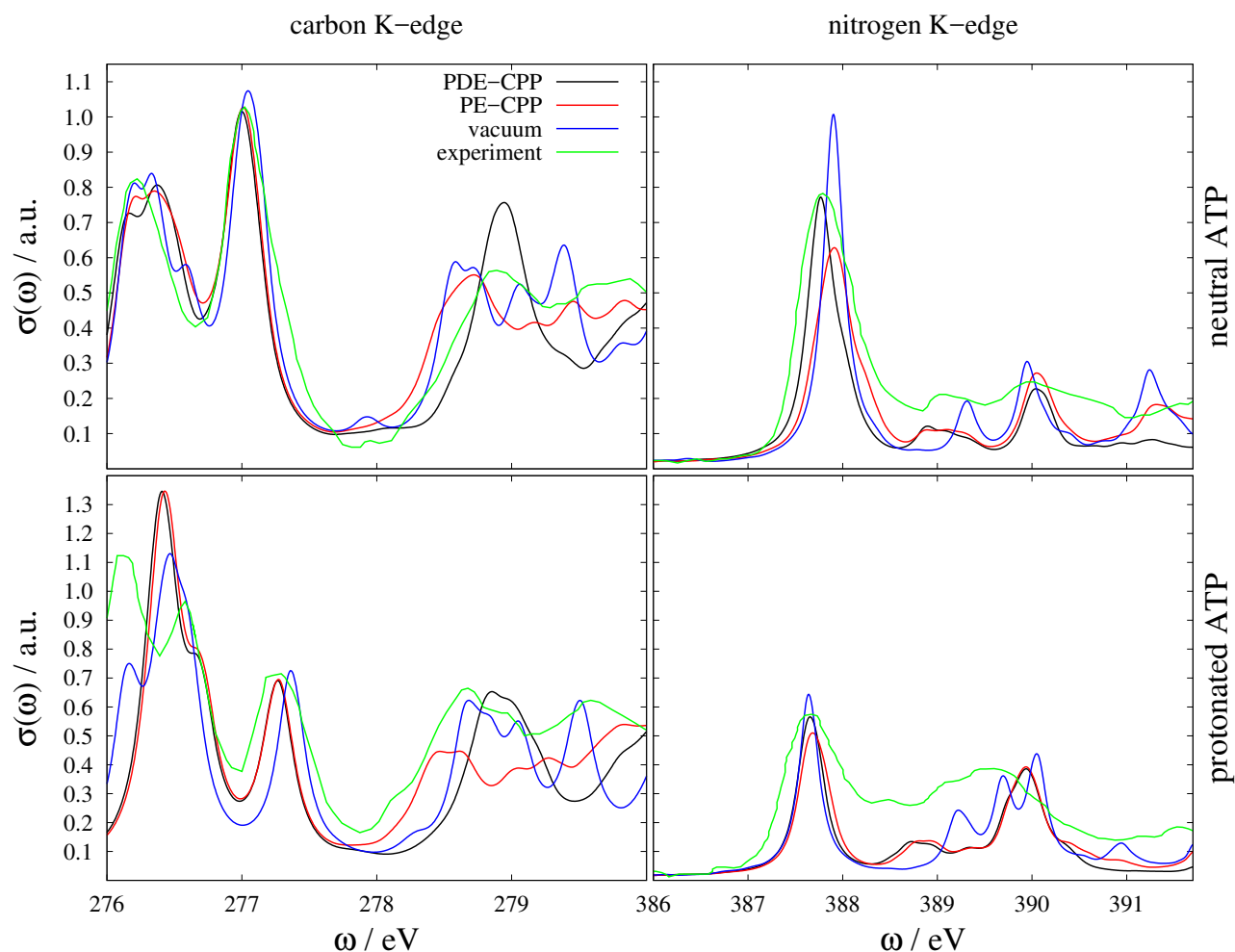


Figure 5: Calculated X-ray spectra of gas phase and aqueous neutral or protonated ATP. The experimental results (aqueous solutions at pH 7.5 and 2.5) from Ref. 66 are also shown. The calculated spectra of the solvated species were obtained as averages based on 10 selected snapshots, see text for details. The experimental results are shifted (by -10.25 eV and -10.4 eV for carbon K-edge and by -12.45 eV and -12.2 eV for nitrogen K-edge) and scaled to be comparable to the calculated results and they were digitized using WebPlotDigitizer.⁶⁰

in the simulated PDE and PE spectra displays a splitting at the top. The analysis of the spectra for each of the 10 snapshots (given in the SI, Fig. S1) indicates that the splitting arises because there are somewhat large spectral differences between the 10 configurations in that region of the spectrum. The analysis given in Table S1 also shows that the peak at 276.4 eV is mostly produced through excitation from the $1s$ -orbital at the **C8** atom, although it also contains contributions from the excitation of atoms **C2** and **C5**. However, each of these contributions is not present for every configuration. The peak at 277.0 eV can, on the contrary, be unambiguously characterized because it originates from the excitation at the **C6** atom for each of the 10 snapshots. There is another set of peaks at around 278 eV–280 eV that arise from excitations of the $1s$ -electrons in the ribose group. We observe a significant difference between PDE and PE in this region, as PDE has a pronounced peak and PE displays only a series of weaker peaks. Such difference has not been observed in the region of absorption of the adenine carbons, possibly because the ribose group contains oxygen atoms, so the virtual orbitals in this region are expected to be more strongly affected by hydrogen bonding with the solvent molecules and these effects could be reflected in excitations from the carbon atoms. The splitting of the intense peak at 286.5 eV experimentally observed in Ref. 66 upon protonation of the **N1** atom at $pH < 4.0$ is not clearly reproduced in our embedded calculations for the corresponding **C2/C5/C8** peak at 276.4 eV. As in the case of adenine, for ATP there also appears a shoulder on the right-hand side of the peak, where the principal contribution to the absorption cross-section is from a **C4** excitation. When looking at the spectra of the snapshots (also in the SI, Fig. S2), we observe that there is very little difference between the spectra based on the different snapshots for the protonated form meaning that both intense peaks can be unambiguously characterized. The reason for this is that there are less configurational fluctuations between the snapshots of the protonated ATP as compared to the neutral ATP because the oxygen atoms of the triphosphate tail interact with the protonated aromatic system in the adenine part, as seen in Fig. S3. Fig. S4 shows the normalized root-mean-square deviation (NRMSD) of $\text{Im} \{ \bar{\alpha}(\omega) \}$ as a function

of the number of configurations taken into account. The NRMSD applied here is defined in eq. S1 in the SI. The NRMSD values are significantly higher for the neutral ATP than for the protonated ATP in the first 2 – 3 snapshots. However, the NRMSD of the neutral form converges more quickly so the NRMSD of both the neutral and protonated forms reaches similar values for the 10 configurations. Fig. S4 also indicates that a satisfying spectral convergence can be achieved even for a relatively small configurational sampling.

At the nitrogen K-edge, the intense peak at 387.8 eV arises from the atoms **N3**, **N7** and **N1**. For PDE this peak is more intense and at a lower frequency than for PE, as is the case for aqueous adenine but opposite from formamide and glycine because the **N3** and **N7** atoms in both adenine and ATP are hydrogen receptors when H-bonding with the nearest water molecules. In the experiment, the intense peak has been found at 400.3 eV. The following peak is a **N10** transition. The position of this peak fluctuates significantly throughout the different snapshots, which results in broadening of the peak. The **N10** peak corresponds to the weak experimental peak at 401.5 eV. The peak at 390.0 eV is a **N9** excitation and it corresponds with the experimental peak at 402.5 eV. The second **N10** peak can only be identified for the PE solvent model at 391.3 eV, while for PDE it is attenuated. The explanation for this is the same as for adenine, namely that it is a probably Rydberg transition, so PDE destabilizes it more strongly through the non-electrostatic repulsion. The PDE result is for this peak more in agreement with the experiment by Kelly et al., which also does not identify this peak for the aqueous ATP. In the protonated spectrum, the leftmost peak is a mixture of **N7** and **N3** excitations for PDE, and a more pure **N7** excitation for PE. The third peak stems from the 1s-excitation at the **N1** and **N9** atoms for both embedding models, and not purely **N9**, as the neutral case. The differences between PDE and PE are more pronounced at the nitrogen K-edge, than at the carbon K-edge (except for the region above 278 eV of the carbon K-edge of ATP).

Conclusion

In this work we have formulated the complex polarization propagator method in combination with polarizable density embedding – PDE-CPP – and applied this model to calculations of the NEXAFS spectra of molecules in solutions. Based on PDE-CPP and PE-CPP, the NEXAFS of aqueous adenine and ATP in both their neutral and protonated forms were simulated and compared to experimental results. In general, both PDE and PE provided reasonably good agreement with the experiments, such as the trend of observing three peaks in the carbon K-edge spectrum of adenine, versus only two peaks at the carbon K-edge of ATP. A significant difference between PDE and PE could be seen for ATP in the region where the $1s$ -electrons from the ribose carbons are being excited and in the case of the Rydberg transition in the nitrogen K-edge spectrum. The cases of formamide and glycine illustrate that the $C1s \rightarrow \pi^*$ transition is generally equally well described with the two embedding models, whereas the differences can be found in the region where Rydberg transitions are expected. For nitrogen and oxygen K-edges the differences are more pronounced because of the different description of hydrogen bonds. From our analysis we can conclude that the X-ray absorption spectra are not significantly affected by the more elaborate description of electrostatic effects through fragment densities and by the inclusion of the non-electrostatic repulsion in PDE versus only multipole expansion in PE, except in the case of transitions involving Rydberg states.

Acknowledgement

Computational resources were provided by the DeIC National HPC Center at the University of Southern Denmark through an Abacus 2.0 grant. J.K. acknowledges financial support from the Villum Foundation and the Danish Council for Independent Research (DFF). S.C. acknowledges support from DTU Chemistry and from the Independent Research Fund Denmark. Support from H2020-MSCA-ITN-2017 training network “COSINE – Computational

Spectroscopy In Natural sciences and Engineering” is also acknowledged.

Supporting Information Available

This material is available free of charge via the Internet at <http://pubs.acs.org/>.

References

- (1) Mennucci, B. Modeling environment effects on spectroscopies through QM/classical models. *Phys. Chem. Chem. Phys.* **2013**, *15*, 6583.
- (2) Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249.
- (3) Lin, H.; Truhlar, D. G. QM/MM: what have we learned, where are we, and where do we go from here? *Theor. Chem. Acc.* **2006**, *117*, 185–199.
- (4) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem. Int. Ed.* **2009**, *48*, 1198–1229.
- (5) Tu, Y.; Laaksonen, A. Implementing Quantum Mechanics into Molecular Mechanics—Combined QM/MM Modeling Methods. *Adv. Quantum Chem.* **2010**, *59*, 1–15.
- (6) van der Kamp, M. W.; Mulholland, A. J. Combined Quantum Mechanics/Molecular Mechanics (QM/MM) Methods in Computational Enzymology. *Biochem.* **2013**, *52*, 2708–2728.
- (7) Severo Pereira Gomes, A.; Jacob, C. R. Quantum-chemical embedding methods for treating local electronic excitations in complex chemical systems. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* **2012**, *108*, 222–277.

- (8) Olsen, J. M. H.; Aidas, K.; Kongsted, J. Excited States in Solution through Polarizable Embedding. *J. Chem. Theory Comput.* **2010**, *6*, 3721–3734.
- (9) Olsen, J. M. H.; Kongsted, J. Molecular properties through polarizable embedding. *Adv. Quant. Chem.* **2011**, *61*, 107–143.
- (10) Olsen, J. M. H.; Steinmann, C.; Ruud, K.; Kongsted, J. Polarizable Density Embedding: A New QM/QM/MM-Based Computational Strategy. *J. Phys. Chem. A* **2015**, *119*, 5344–5355.
- (11) Olsen, J.; Jørgensen, P. Linear and nonlinear response functions for an exact state and for an MCSCF state. *J. Chem. Phys.* **1985**, *82*, 3235–3264.
- (12) Salek, P.; Vahtras, O.; Helgaker, T.; Ågren, H. Density-functional theory of linear and nonlinear time-dependent molecular properties. *J. Chem. Phys.* **2002**, *117*, 9630–9645.
- (13) Christiansen, O.; Jørgensen, P.; Hättig, C. Response functions from Fourier component variational perturbation theory applied to a time-averaged quasienergy. *Int. J. Quant. Chem.* **1998**, *68*, 1–52.
- (14) Helgaker, T.; Coriani, S.; Jørgensen, P.; Kristensen, K.; Olsen, J.; Ruud, K. Recent Advances in Wave Function-Based Methods of Molecular-Property Calculations. *Chem. Rev.* **2012**, *112*, 543–631.
- (15) Norman, P.; Bishop, D. M.; Jensen, H. J. A.; Oddershede, J. Near-resonant absorption in the time-dependent self-consistent field and multiconfigurational self-consistent field approximations. *J. Chem. Phys.* **2001**, *115*, 10323–10334.
- (16) Norman, P.; Bishop, D. M.; Jensen, H. J. A.; Oddershede, J. Nonlinear response theory with relaxation: The first-order hyperpolarizability. *J. Chem. Phys.* **2005**, *123*, 194103.

- (17) Norman, P. A perspective on nonresonant and resonant electronic response theory for time-dependent molecular properties. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20519–20535.
- (18) Kristensen, K.; Kauczor, J.; Kjærgaard, T.; Jørgensen, P. Quasienergy formulation of damped response theory. *J. Chem. Phys.* **2009**, *131*, 044112.
- (19) Kauczor, J.; Jørgensen, P.; Norman, P. On the Efficiency of Algorithms for Solving Hartree–Fock and Kohn–Sham Response Equations. *J. Chem. Theory Comput.* **2011**, *7*, 1610–1630.
- (20) Pedersen, M. N.; Hedegård, E. D.; Olsen, J. M. H.; Kauczor, J.; Norman, P.; Kongsted, J. Damped Response Theory in Combination with Polarizable Environments: The Polarizable Embedding Complex Polarization Propagator Method. *J. Chem. Theory Comput.* **2014**, *10*, 1164–1171.
- (21) List, N. H.; Coriani, S.; Kongsted, J.; Christiansen, O. Lanczos-driven coupled-cluster damped linear response theory for molecules in polarizable environments. *J. Chem. Phys.* **2014**, *141*, 244107.
- (22) Schwabe, T.; Olsen, J. M. H.; Sneskov, K.; Kongsted, J.; Christiansen, O. Solvation Effects on Electronic Transitions: Exploring the Performance of Advanced Solvent Potentials in Polarizable Embedding Calculations. *J. Chem. Theory Comput.* **2011**, *7*, 2209–2217.
- (23) Fradelos, G.; Wesolowski, T. A. Importance of the Intermolecular Pauli Repulsion in Embedding Calculations for Molecular Properties: The Case of Excitation Energies for a Chromophore in Hydrogen-Bonded Environments. *J. Phys. Chem. A* **2011**, *115*, 10018–10026.
- (24) Reinholdt, P.; Kongsted, J.; Olsen, J. M. H. Polarizable Density Embedding: A Solution

- to the Electron Spill-Out Problem in Multiscale Modeling. *J. Phys. Chem. Lett.* **2017**, 5949–5958.
- (25) Nåbo, L. J.; Olsen, J. M. H.; List, N. H.; Solanko, L. M.; Wüstner, D.; Kongsted, J. Embedding beyond electrostatics—The role of wave function confinement. *J. Chem. Phys.* **2016**, *145*, 104102.
- (26) Lindon, J. C. In *Encyclopedia of Spectroscopy and Spectrometry (Second Edition)*, second edition ed.; Lindon, J. C., Ed.; Academic Press: Oxford, 2010; pp 2058 – 2068.
- (27) Smith, J. W.; Saykally, R. J. Soft X-ray Absorption Spectroscopy of Liquids and Solutions. *Chem. Rev.* **2017**, DOI: 10.1021/acs.chemrev.7b00213.
- (28) Huzinaga, S.; Cantu, A. A. Theory of Separability of Many-Electron Systems. *J. Chem. Phys.* **1971**, *55*, 5543–5549.
- (29) Olsen, J. M. H. Development of Quantum Chemical Methods towards Rationalization and Optimal Design of Photoactive Proteins. Ph.D. thesis, University of Southern Denmark, Odense, Denmark, 2012; DOI: 10.6084/m9.figshare.156852.
- (30) Raymond, X. S. *Elementary introduction to the theory of pseudodifferential operators*; CRC Press: Boca Raton, Florida, 1991.
- (31) Applequist, J.; Carl, J. R.; Fung, K.-K. Atom dipole interaction model for molecular polarizability. Application to polyatomic molecules and determination of atom polarizabilities. *J. Am. Chem. Soc.* **1972**, *94*, 2952–2960.
- (32) Sneskov, K.; Schwabe, T.; Kongsted, J.; Christiansen, O. The polarizable embedding coupled cluster method. *J. Chem. Phys.* **2011**, *134*, 104108.
- (33) MarvinSketch (version 17.26.0), ChemAxon Ltd. (<http://www.chemaxon.com>), 2017.
- (34) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and

- transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
- (35) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. Efficient diffuse function-augmented basis sets for anion calculations. III. The 3-21+G basis set for first-row elements, Li–F. *J. Comput. Chem.* **1983**, *4*, 294–301.
- (36) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements. *J. Chem. Phys.* **1982**, *77*, 3654–3665.
- (37) Hariharan, P. C.; Pople, J. A. The influence of polarization functions on molecular orbital hydrogenation energies. *Theor. Chem. Acc.* **1973**, *28*, 213–222.
- (38) Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self—Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian—Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56*, 2257–2261.
- (39) Frisch, M. J. et al. Gaussian09 Revision D.01. Gaussian Inc. Wallingford CT 2009.
- (40) Schrödinger release 2017-2: Maestro, LLC, New York, NY. 2017.
- (41) Schrödinger Release 2017-2: Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, 2017. Maestro-Desmond Interoperability Tools, Schrödinger, New York, NY. 2017.
- (42) Hršak, D.; Olsen, J. M. H.; Kongsted, J. Optimization and transferability of non-electrostatic repulsion in the polarizable density embedding model. *J. Comp. Chem.* **2017**, *38*, 2108–2117.
- (43) Banks, J. L. et al. Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J. Comput. Chem.* **2005**, *26*, 1752–1780.

- (44) QSite 2016-3, Schrödinger, LLC, New York, NY. 2016.
- (45) Murphy, R. B.; Philipp, D. M.; Friesner, R. A. A mixed quantum mechanics/molecular mechanics (QM/MM) method for large-scale modeling of chemistry in protein environments. *J. Comp. Chem.* **2000**, *21*, 1442–1457.
- (46) Bochevarov, A. D.; Harder, E.; Hughes, T. F.; Greenwood, J. R.; Braden, D. A.; Philipp, D. M.; Rinaldo, D.; Halls, M. D.; Zhang, J.; Friesner, R. A. Jaguar: A high-performance quantum chemistry software program with strengths in life and materials sciences. *Int. J. Quant. Chem.* **2013**, *113*, 2110–2142.
- (47) Sneskov, K.; Matito, E.; Kongsted, J.; Christiansen, O. Approximate Inclusion of Triple Excitations in Combined Coupled Cluster/Molecular Mechanics: Calculations of Electronic Excitation Energies in Solution for Acrolein, Water, Formamide, and N-Methylacetamide. *J. Chem. Theory Comput.* **2010**, *6*, 839–850.
- (48) Aidas, K.; Kongsted, J.; Sabin, J. R.; Oddershede, J.; Mikkelsen, K. V.; Sauer, S. P. A. The Effect of Solvation on the Mean Excitation Energy of Glycine. *J. Phys. Chem. Lett.* **2010**, *1*, 242–245.
- (49) Aquilante, F. et al. Molcas 8: New capabilities for multiconfigurational quantum chemical calculations across the periodic table. *J. Comp. Chem.* **2016**, *37*, 506–541.
- (50) Gagliardi, L.; Lindh, R.; Karlström, G. Local properties of quantum chemical systems: The LoProp approach. *J. Chem. Phys.* **2004**, *121*, 4494–4500.
- (51) Yanai, T.; Tew, D. P.; Handy, N. C. A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51–57.
- (52) Ekström, U.; Norman, P. X-ray absorption spectra from the resonant-convergent first-order polarization propagator approach. *Phys. Rev. A* **2006**, *74*, 042722.

- (53) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (54) Woon, D. E.; Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. *J. Chem. Phys.* **1993**, *98*, 1358–1371.
- (55) Balabanov, N. B.; Peterson, K. A. Systematically convergent basis sets for transition metals. I. All-electron correlation consistent basis sets for the 3d elements Sc–Zn. *J. Chem. Phys.* **2005**, *123*, 064107.
- (56) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
- (57) Aidas, K. et al. The Dalton quantum chemistry program system. *WIREs: Comput. Mol. Sci.* **2014**, *4*, 269–284.
- (58) Fransson, T.; Saue, T.; Norman, P. Four-Component Damped Density Functional Response Theory Study of UV/Vis Absorption Spectra and Phosphorescence Parameters of Group 12 Metal-Substituted Porphyrins. *J. Chem. Theory Comput.* **2016**, *12*, 2324–2334.
- (59) Plekan, O.; Feyer, V.; Richter, R.; Coreno, M.; de Simone, M.; Prince, K.; Trofimov, A.; Gromov, E.; Zaytseva, I.; Schirmer, J. A theoretical and experimental study of the near edge X-ray absorption fine structure (NEXAFS) and X-ray photoelectron spectra (XPS) of nucleobases: Thymine and adenine. *Chem. Phys.* **2008**, *347*, 360–375.
- (60) Rohatgi, A. WebPlotDigitizer, version 4.0, <https://automeris.io/WebPlotDigitizer>. 2017.
- (61) Besley, N. A.; Hirst, J. D. Ab Initio Study of the Effect of Solvation on the Electronic

- Spectra of Formamide and N-Methylacetamide. *J. Phys. Chem. A* **1998**, *102*, 10791–10797.
- (62) Besley, N. A.; Hirst, J. D. Ab Initio Study of the Electronic Spectrum of Formamide with Explicit Solvent. *J. Am. Chem. Soc.* **1999**, *121*, 8559–8566.
- (63) Ottosson, N.; Aziz, E. F.; Bergersen, H.; Pokapanich, W.; Öhrwall, G.; Svensson, S.; Eberhardt, W.; Björneholm, O. Electronic Rearrangement upon the Hydrolyzation of Aqueous Formaldehyde Studied by Core–Electron Spectroscopies. *J. Phys. Chem. B* **2008**, *112*, 16642–16646.
- (64) Ishii, I.; Hitchcock, A. P. A quantitative experimental study of the core excited electronic states of formamide, formic acid, and formyl fluoride. *J. Chem. Phys.* **1987**, *87*, 830–839.
- (65) Messer, B. M.; Cappa, C. D.; Smith, J. D.; Wilson, K. R.; Gilles, M. K.; Cohen, R. C.; Saykally, R. J. pH Dependence of the Electronic Structure of Glycine. *J. Phys. Chem. B* **2005**, *109*, 5375–5382.
- (66) Kelly, D. N.; Schwartz, C. P.; Uejio, J. S.; Duffin, A. M.; England, A. H.; Saykally, R. J. Communication: Near edge x-ray absorption fine structure spectroscopy of aqueous adenosine triphosphate at the carbon and nitrogen K-edges. *J. Chem. Phys.* **2010**, *133*, 101103.

Graphical TOC Entry

